DOCUMENT RESUME

ED 318 761 TM 014 808

AUTHOR Hand, Carol A.; Prather, James E.
TITLE Investigating High School Effectiveness Using College Performance: An Analysis of Residuals.
PUB DATE Apr 90
NOTE 14p.; Paper presented at the Annual Meeting of the American Educational Research Association (Boston, MA, April 16-20, 1990).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academic Achievement; Academic Aptitude; *College Freshmen; Educational Quality; Effective Schools Research; Grade Point Average; Higher Education; High School Graduates; High Schools; *Performance Factors; Predictive Measurement; *Predictor Variables; Regression (Statistics); *School Effectiveness; School Role
IDENTIFIERS Scholastic Aptitude Test

ABSTRACT

The purpose of this study was to relate college-level or univeristy-level student performance to high school effectiveness, which was measured in terms of the college performance of a high school's graduates, after controlling for students' academic aptitude. Unusually effective schools were identified by examining the residuals of a multiple regression analysis. It was hypothesized that certain schools would be unusually effective, in that graduates would have higher grade point averages (GPAs) than would be expected given their academic aptitude as measured by the Scholastic Aptitude Test (SAT) and by their socioeconomic environment. The analysis was conducted using 111,703 freshmen in a large univesity system from 301 high schools and 162 school systems. Multiple regression equations predicting college GPA were developed for each institution by gender and minority status. Independent variables included: SAT scores; the proportions of free and reduced lunches at each high school; and college experience as measured by credit hours attempted and earned. There was a clear tendency for rural schools to nurture graduates who achieved higher GPAs than SAT scores or socioeconomic status would predict. Urban schools were more likely to produce students of lower achievement than predicted. The method appears promising for extending assessment perspectives beyond internal college or university concerns. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

CAROL HAND

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED318761

INVESTIGATING HIGH SCHOOL EFFECTIVENESS USING COLLEGE PERFORMANCE: AN ANALYSIS OF RESIDUALS

Carol A. Hand

and

James E. Prather

Office of Institutional Research
Georgia State University
Atlanta, Georgia 30303
(404) 651-2570

Paper presented at the 1990
Annual Meeting
of the American Educational Research Association
Boston, Massachusetts
April 16-20, 1990

TM 014808

Investigating High School Effectiveness Using College Performance: An Analysis of Residuals

As assessment programs have been implemented, the issue of how to integrate secondary-level and college-level measures of effectiveness remains. Exploratory research is needed in this section of the assessment field.

The purpose of this paper is to relate college-level or university-level student performance with high school effectiveness. A school's effectiveness is measured in terms of the college performance of its graduates, after controlling for students' academic aptitude. The method of this study is to identify unusually effective schools by examining the residuals of a multiple regression analysis. It is hypothesized that certain schools would be identified as unusually effective, in that graduates of these schools have higher college grade point averages than would be expected given their academic aptitude as measured by the Scholastic Aptitude Test (SAT) and by socioeconomic environment.

Relevant Previous Research

In the 1960s and 1970s educational researchers attempted to investigate educational effectiveness. Several large-scale statistical studies failed to find consistent relationships between school differences and students' performances on achievement tests (Averch et al., 1972; Coleman, 1966; Jencks et al., 1972).

In spite of the inconclusive or negative results of many previous school effectiveness studies, it is possible that some

unusually effective schools do exist. The problem is one of identifying those schools that seem to do an exceptional job in terms of student academic performance. Schools which may be exceptional in terms of their incremental effectiveness can then be identified by finding the outliers in the residuals of the regression analysis. This method has been successfully used by Klitgaard and Hall (1977), Austin (1979), Clauset and Gaynor (1981), and Edmonds (1979), and is recommended as potentially useful by Anderson (1982), Levine (1973) and Klitgaard (1978).

Anderson (1982) provides an in-depth review of school climate and school effects research. In a discussion of input variables, the necessity of controlling for variations in student inputs is stressed. The increment in achievement that the school provides is seen as the appropriate measure of a school's quality. The author considers the inconclusive findings of many school effect studies to be a result of design and analysis problems, including inadequacy of statistics, misinterpretation of statistical results, misuse of statistical procedures and inadequacy of cross-sectional designs. The author recommends several design alternatives, including the use of outliers as described by Klitgaard and Hall (1977) and others.

Levine (1973) provides an extensive review and criticism of attempts by Coleman (1966), Jencks et al. (1972) and others to investigate educational effectiveness. Problems of data that characterize many input-output studies include: aggregation of data at the school or district level; the inadequacy of standardized

achievement tests as a summary measure of school effectiveness; and the diminished chances of discovering significant input-output relationships by confining the analysis to average school effects. Several recommendations for future research are made, including analysis of the residuals from multiple regressions to identify unusual schools.

In a large-scale study by Walberg and Fowler (1987), district socioeconomic status, per-student expenditures on education and size of enrollment by district were used to predict average standardized test scores at three grade levels. The authors also controlled only for socioeconomic status and examined the residuals for districts high or low on achievement measures. It was found that the districts with the highest residuals (highest achievement after adjusting for socioeconomic status) expended the least amount of money for education. When socioeconomic status and per-student expenditures were taken into account, smaller districts tended to have higher achievement than larger districts.

The idea for this paper came from a study by Klitgaard and Hall (1977) titled "A Statistical Search for Unusually Effective Schools," which was included in the book, Statistics and Public Policy. The authors began by describing studies that have failed to find consistent relationships between school variables and student achievement, and followed with possible explanations for this failure. They then develop the idea of looking at the outliers in the residuals of multiple regression analysis.

Six sets of data from Michigan and New York City schools were examined. Socioeconomic and other non-school background factors were included in the regression equation, while school variables were explicitly excluded. It was assumed that the variation remaining after such a statistical fit would represent school effectiveness along with random variation (which was expected to be small). It was also decided to avoid overcontrolling for too many background variables in order to give exceptional schools every opportunity to show up.

Histograms of the residuals on socioeconomic factors from a regression of school achievement scores were visually inspected for unusual right tails. Next, the series of distributions of residuals were looked at over many years to see if the same schools were consistently above the mean. More schools with residuals greater than one standard deviation above the mean were found than would have been expected based on chance alone.

The next step was to determine if the top schools (those with high positive residuals) were different from the other schools in terms of any non-school factors that had been excluded from the regression equation. It was decided at this point to run new regressions to include controls for racial composition and community type. The addition of these variables in the regression reduced the total variability and allowed unusually effective schools to stand out, as evidenced by the unusual right tails in histograms of the new residuals. It was found that the top 15 schools were predominantly below the state averages on both

socioeconomic status and minority enrollment and tended to be located in rural areas. This was true in spite of controlling for these variables in the regressions. Possible explanations, including sample sizes and heteroskedasticity, were tested and discussed.

The final task was to compare the top and average schools in terms of means and of standard deviations on several school, background and other variables. Important differences appeared on several school-related variables, while none of the background variables showed significant differences.

In summary, the authors used multiple regression techniques to predict standardized achievement test performance while controlling for various background influences. Residuals from the regression lines were assumed to be a function of school factors and were conceptualized as a measure of school effectiveness. Schools with high positive residuals were compared with average schools and found to differ on several important school variables. The authors conclude that the methodology developed in their report should be useful in future attempts to investigate school effectiveness.

A study by Hand and Prather (1989) using a methodology similar to Klitgaard and Hall (1977) investigated the possibility of using college GPAs as an indicator of unusually effective high schools. They used data from over 150,00 first-time freshmen in a large state college and university system. Their analysis developed multiple regression equations predicting college GPA for each

institution by gender and minority status. The independent variables were SAT verbal and mathematics scores, college experience as measured by credit hours attempted and earned, and the proportion of free and reduced-price lunches at each of their high schools. The predicted college GPA was subtracted from the students' actual GPA to calculate a residual GPA. The average residual GPA for graduates of each high school was determined. As multiple regression diagnostics were performed on these prediction equations, it was found that credit hours carried and earned showed clear evidence of a quadratic relationship with GPA. Thus squared values of credit hours attempted and earned were also included in the prediction equations. These equations are reported in Hand and Prather (1989) and showed that the multiple correlation coefficients squared (R^2) for the institution equations are generally between .40 to .60. Socioeconomic status of the high school had little independent influence on the college GPA.

The regression coefficients for the socioeconomic status indicator for the high schools vary considerably in both magnitude and sign. No patterns seem present and very few of these coefficients are even twice their standard errors. Thus the predictions would be almost identical even if a correction for high school socioeconomic status were not included in the equations. In contrast to the coefficients for free and reduced-price lunches, the coefficients for SAT scores are more consistent, with more than 90 percent having positive signs. In a few cases the coefficients

have negative signs. Also, the coefficient of SAT verbal score is usually larger than that of the SAT mathematics score.

Patterns are evident in the control variables used to account for college experience. The coefficient for credit hours earned is normally positive, while for credit hours carried it is negative. Thus the more credit hours earned the higher the predicted GPA. The squared values show opposite patterns, with coefficient of hours earned squared being negative and hours attempted squared producing positive coefficient in most equations. The residual college GPA was obtained for each individual student by subtracting actual GPA from predicted GPA. These residuals were then averaged across the students' school systems to develop a mean residual GPA for each school system.

Hand and Prather (1989) report that the mean residual GPAs show a pattern that high schools located in urban and suburban areas tend to have negative mean residual GPAs. On the other hand, there are high schools located in rural areas and small towns that have positive mean residual GPAs.

Methods

This paper undertakes a further analysis of the data collected for Hand and Prather (1989) in an effort to measure school system effectiveness. The database was comprised of 144,848 first-time freshmen in a large state university system. The analysis developed multiple regression equations predicting college GPA for each institution by gender and minority status. The independent

variables were SAT verbal and mathematics scores, the proportions of free and reduced lunches at each high school, and college experience as measured by credit hours attempted and earned.

Steps in the analysis are as follows:

1) Multiple regression equations for each system institution were developed by gender and minority status. The level of analysis is the individual student.

2) The dependent variable in the regression equations is the college GPA. The independent variables are SAT verbal and mathematics scores, and credit hours accumulated as a control for college experience. A control measure of the socioeconomic status of the high school--the proportion of students who receive free and reduced-price lunches--is also included in the multiple regression equations.

3) Using the appropriate regression equation for the institution attended, the residual (the difference between actual and predicted GPA) is calculated for each student.

4) The average residual for graduates of each school system is determined. School systems are ranked based on their mean residuals, and high-residual and low-residual systems examined more closely for similarities and differences within and between the groups.

Findings

The research question of whether differences that appeared at the high school level would appear at the local school system level were analyzed using a one-way analysis of variance with nesting design. To do this analysis the residuals were used in a one-way analysis of variance by school system with high school as the nesting factor. The analysis was done with 111,703 students in 162 school systems and 301 high schools. This analysis resulted in an F value of 4.19 with a $p < .001$ (df=288, 111702) and an R^2 of .0107. The school system means for the residual variable and other

variables such as high school average, SAT scores and percent of female and black students are given in Table 1.

The mean residuals show a clear tendency that rural school systems are likely to graduate students who achieve higher GPAs in college than their SAT scores or the socioeconomic status of their high school would indicate. Examples of students from rural area school systems with college-level performance higher than predicted are Fannin (mean residual of .15), Whitfield (.15), Dawson (.13), Franklin (.14), Habersham (.11), Lumpkin (.12), Hogansville (.12), Madison (.15), Oglethorpe (.12), Social Circle (.13), Bleckley (.16), Evans (.15), Vidalia (.11), Warren (.15), Thomas (.15), Bacon (.16) and Clinch (.12). On the other hand, the school systems in urban areas tend to produce students whose performance in college generally falls below predicted levels. School systems in urban locations indicate means residuals of -.04 for Atlanta, DeKalb and Fulton; -.03 for Gwinnett; -.02 for Marietta; and -.01 for Cobb.

While not all rural school systems produce students whose college performance is generally higher than predicted by SAT scores and socioeconomic level of high school, many do. College students from urban school systems are less likely to evidence performance above that predicted by SAT scores and socioeconomic status of their high school. This suggests that some small town and rural school systems do give their students something "extra," so that these students perform at a higher level in college than would be predicted by these measures.

Summary and Conclusions

The objective of this paper has been to relate college-level or university-level student performance with high school effectiveness. This effort is largely exploratory due to the preliminary nature of research in this area of the assessment field. It is hypothesized that college or university GPAs can be used to identify unusually effective school systems in that the graduates of some high schools have higher college GPAs than would be expected given their academic aptitude as measured by SAT scores and given the socioeconomic status of their high school.

This paper focuses upon whether differences that appeared at the high school level would appear at the local school system level. To do this analysis the residuals were used in a one-way analysis of variance by school system with high school as the nesting factor. The analysis was done with 111,703 students in 162 school systems and 301 high schools. This analysis resulted in an F value of 4.19 with a $p < .001$ (df=288, 111702) and an R^2 of .0107.

The mean residuals show a clear tendency that rural school systems are likely to graduate students who achieve higher GPAs in college than their SAT scores or the socioeconomic status of their high school would indicate. On the other hand, the school systems in urban areas tend to produce students whose performance in college generally falls below predicted levels.

It is felt that this method has promise in extending assessment perspectives and possibilities beyond internal college or

university concerns. While caution must be exercised, the uniting of levels of instruction does seem to be a forward step in assessment efforts.

Implications

Substantive and methodological implications are supported by this paper. Finding a linkage between measures of effectiveness on the school system and college levels can be an important substantive contribution to the assessment field. It is felt that this is of value to those concerned with methods of identifying effective systems schools, to those interested in the relationship between school system attended and college GPA and to those interested in school system performance as an indicator of college potential.

References

Anderson, C.S. (1982). The search for school climate: A review of the research. Review of Educational Research, 52, 368-420.

Austin, G.R. (1979). Exemplary schools and the search for effectiveness. Educational Leadership, 37, 10-12; 14.

Averch, H.S., et al. (1972). How effective is schooling? A critical review and synthesis of research findings. Santa Monica, CA: The Rand Corporation.

Clauset, K.H., Jr., & Gaynor, A.K. (1981). The dynamics of effective and ineffective schooling: A preliminary report of a system dynamics policy study. Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, April 1981.

Coleman, J.S., et al. (1966). Equality of educational opportunity. Washington, DC: U.S. GPO.

Edmonds, R.R. (1979). Effective Schools for the Urban Poor. Educational Leadership, 37, 15-18; 20-24.

Hand, C.A. & Prather, J.E. (1989). Using college grade point average in assessment research. Paper presented at the annual Conference of the Southern Association for Institutional Research, Durham, North Carolina, October 1989.

Jencks, C.S. et al. (1972). Inequality. New York: Basic Books.

Klitgaard, R.E. (1978). Identifying exceptional performers. Policy Analysis, 4, 529-547.

Klitgaard, R.E. & Hall, G.R. (1977). A statistical search for unusually effective schools. In W.W. Fairley & F. Mosteller (Eds.), Statistics and Public Policy, (pp.51-86). Reading: MA: Addison-Wesley.

Levine, D.M. (1973). Educational policy after inequality. Teachers College Record, 75, 149-179.

Walberg, H.J., & Fowler, W.J., Jr. (1987). Expenditure and size efficiencies of public school districts. Educational Researcher, 16, 5-13.